

## Tilburg University

### **Verbal suggestions about treatment effectiveness do not modulate the effectiveness of a laboratory model of EMDR therapy**

Mertens, Gaëtan; Schie, Kevin van; Lammertink, Sophie; Littel, Marianne; Engelhard, Iris

DOI:

[10.31234/osf.io/zhcnp](https://doi.org/10.31234/osf.io/zhcnp)

Publication date:

2020

Document Version

Early version, also known as pre-print

[Link to publication in Tilburg University Research Portal](#)

*Citation for published version (APA):*

Mertens, G., Schie, K. V., Lammertink, S., Littel, M., & Engelhard, I. (2020). *Verbal suggestions about treatment effectiveness do not modulate the effectiveness of a laboratory model of EMDR therapy: Results of two preregistered studies*. PsyArXiv Preprints. <https://doi.org/10.31234/osf.io/zhcnp>

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**Verbal suggestions about treatment effectiveness do not modulate the effectiveness of a laboratory model of EMDR therapy: Results of two preregistered studies**

Word count (excluding tables and footnotes): 5968

Gaëtan Mertens<sup>1, 2</sup>, Kevin van Schie<sup>3, 4</sup>, Sophie Lammertink<sup>1</sup>, Marianne Littel<sup>3</sup>, and Iris M. Engelhard<sup>1</sup>

<sup>1</sup>Department of Clinical Psychology, Utrecht University, Utrecht, the Netherlands

<sup>2</sup> Department of Medical and Clinical Psychology, Tilburg University, Tilburg, the Netherlands

<sup>3</sup>Psychology, Education & Child Studies, Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam, Rotterdam, the Netherlands

<sup>4</sup>MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK

Correspondence concerning this article should be addressed to Gaëtan Mertens, Department of Medical and Clinical Psychology, Warandelaan 2, room T526, Tilburg University, 5037 AB Tilburg, the Netherlands. E-mail: [g.mertens@tilburguniversity.edu](mailto:g.mertens@tilburguniversity.edu)

**Funding details:** The research reported in this paper was funded by a NWO Vici grant (grant number: 453-15-005) awarded to Iris M. Engelhard.

**Data availability statement:** The data of the experiments reported in this article are available at <https://osf.io/mu3ca/> and <https://osf.io/m7crd/>.

**Disclosure Statement:** The authors declare no conflict of interest regarding the research reported in this article.

### **Abstract**

**Background:** As for many psychotherapies, there is ongoing discussion about the role of specific versus unspecific mechanisms in the effectiveness of Eye-Movement and Desensitization Reprocessing (EMDR) therapy for post-traumatic stress disorder. However, research directly examining the potential role of non-specific mechanisms in EMDR is scarce.

**Objectives:** Here, we address the role of one non-specific factor that is often put forward, namely treatment effectiveness expectations, in a laboratory model of EMDR therapy.

**Method:** In a lab-based study ( $N = 96$ ) and an online study ( $N = 173$ ), we gave participants verbal instructions to manipulate their treatment expectations. Instructions emphasized EMDR's effectiveness or ineffectiveness. Then, participants were asked to recollect an unpleasant autobiographical memory with or without making eye-movements.

**Results:** In line with previous studies, we found significant reductions of reported vividness and emotionality of negative autobiographical memories. However, this effect was not modulated by manipulated treatment efficacy expectations. It must be noted that manipulation checks indicated that it was difficult to influence treatment effectiveness expectations using verbal suggestions.

**Conclusions:** These findings corroborate the results of two earlier reports, suggesting that the role of treatment expectations in EMDR therapy may be limited.

**Keywords:** EMDR therapy; working memory taxation; Memory degradation; Non-specific treatment factors; Expectation effects

**Highlights**

- Outcome expectations are thought to partly explain psychotherapy effectiveness
- Few studies have addressed this for EMDR therapy
- We report result of 2 laboratory analogue studies (N = 96; N = 173)
- Intervention effects were not moderated by treatment effectiveness suggestions

**Verbal suggestions about treatment effectiveness do not modulate the effectiveness of a laboratory model of EMDR therapy: Results of two preregistered studies**

Eye-Movement Desensitisation and Reprocessing (EMDR) therapy (Shapiro, 2017) is a well researched and effective therapy for the treatment of Post-Traumatic Stress Disorder (PTSD) (Lewis et al., 2020). EMDR therapy comprises several interventions, but the main component is that the patient is asked to recall an aversive memory while simultaneously making horizontal eye-movements (Shapiro, 2017). Lab studies have shown that this component reduces the emotionality and vividness of emotional memories (Andrade et al., 1997; Engelhard et al., 2019).

Laboratory research suggests that a working mechanism of EMDR therapy involves taxing of working memory (WM) (Andrade et al., 1997; Gunter & Bodner, 2008; van den Hout & Engelhard, 2012). According to WM theories, WM is a limited capacity system used for the storage, manipulation and retrieval of mental images and memories (Baddeley, 2012). Making eye-movements taxes WM, which leaves less capacity for the retrieval of an emotional memory – a process that also requires WM capacity (Andrade et al., 1997). It has been hypothesized that the degraded emotional memory will be less emotionally intense, and will then either be restored in a degraded fashion (van den Hout & Engelhard, 2012) or reappraised (Engelhard et al., 2019; Gunter & Bodner, 2008), resulting in long-term decreases in memory vividness and emotionality and, in the context of therapy, a reduction of symptoms (Gunter & Bodner, 2008).

As with many interventions, the question arises whether part of EMDR therapy's effectiveness can be ascribed to non-specific factors, such as expectation effects (e.g., Devilly, 2005; Lilienfeld, 1996). Outcome expectations can be defined as prognostic beliefs about the (beneficial) consequences of engaging in treatment (Constantino et al., 2011). Such expectations

seem to contribute substantially to the effectiveness of psychotherapies, such as cognitive-behavioural therapy, interpersonal therapy, and psychodynamic therapy (e.g., Arnkoff et al., 2002; Constantino et al., 2011; Greenberg et al., 2006). For example, in a meta-analysis of 46 studies, Constantino et al. (2011) found a weighted effect size of Cohen's  $d = .24$  of psychotherapy outcome expectations on post treatment symptom reduction. Yet, patients' expectations are often neglected in psychotherapy research, despite the fact that adequately informing patients and managing expectations is indispensable in clinical practice (Arnkoff et al., 2002; Greenberg et al., 2006). Indeed, also in EMDR therapy, setting expectations is part of the treatment protocol (Shapiro, 2017).

To our knowledge, only Gosselin and Matthews (1995) and Littel, van Schie, and van den Hout (2017; two studies) have investigated the effects of treatment expectations in a laboratory analogue of EMDR therapy. Both reports found that treatment expectations about EMDR therapy did not seem to moderate the effectiveness of the EMDR analogue intervention. However, several limitations to these studies should be taken into account. First, both used relatively small sample sizes (each approximately 20 participants per condition), which did not provide adequate statistical power to detect medium- and small-sized effects of expectation effects. Second, the first experiment in Littel et al. (2017) selected participants with prior knowledge, but did not provide information about the effectiveness of the eye-movements intervention in EMDR therapy. Moreover, in their second experiment, the instructions were focused on the hypothetical working mechanism, rather than EMDR's effectiveness per se. Similarly, Gosselin and Matthews (1995) merely informed participants in the low-expectation condition that little was known about the effectiveness EMDR. It could be argued that these manipulations provided limited information about effectiveness of EMDR therapy, which differs from clinical practice in which

patients receive more explicit information either from their therapist or by looking up information themselves. Hence, it is important to gain insight into the effects of more explicit instructions about its effectiveness. Third, these earlier studies did not include a follow-up test, even though previous studies have shown that effectiveness of the eye-movement intervention tends to decrease over time (e.g., van Veen et al., 2020).

Taking these earlier limitations into account, the current research addresses whether inducing either positive or negative expectations about the effectiveness of EMDR therapy influences the effectiveness of a laboratory analogue of EMDR therapy. In line with the effect of treatment effectiveness expectations in other psychotherapies (e.g., Constantino et al., 2011), we expected that inducing positive expectations would lead to increased effectiveness of the laboratory analogue of EMDR therapy (i.e., the eye-movements intervention) and that inducing negative expectations would lead to decreased effectiveness of this intervention.

## **Experiment 1**

### **Pre-registration**

The sample size determination, design, procedure, and data analyses steps were pre-registered on the Open Science Framework prior to the data collection (<https://osf.io/m7crd/>).

### **Participants**

In order to improve the statistical power of this study compared to the prior studies (Gosselin & Matthews, 1995; Littel et al., 2017), the sample size in each of the conditions of this experiment was increased with approximately 50% to 32 participants per condition (96 participants in total across the three conditions, see below). Power calculations using G-Power indicated a statistical power of .13, .57, and .94 to detect small ( $f = .10$ ), medium ( $f = .25$ ), and large ( $f = .40$ ) effects, respectively (Faul et al., 2007). Participants were recruited via posters on

the Utrecht University campus and Facebook groups. We selected participants according to their answers on a pre-screening questionnaire, which was completed by 202 persons. Eventually, 106 persons were excluded from participation based on the following self-report criteria: bad eyesight, interfering neurological or medical conditions, psychological complaints, currently receiving psychological treatment, having experienced traumatic experiences which they still suffered from, and/or having participated in studies or treatment comparable to our procedure. This led to a final sample size of 96 (see Table 1). All participants provided written informed consent and received partial credit towards a course requirement or financial reimbursement (€8). Participants were randomly assigned to one of the three different conditions in this first experiment.

Table 1.

*Descriptive statistics of the different conditions in Experiment 1.*

|                          | Positive information ( $n = 32$ ) | Negative information ( $n = 32$ ) | Neutral information ( $n = 32$ ) |
|--------------------------|-----------------------------------|-----------------------------------|----------------------------------|
| Sex (men/women)          | 5/27                              | 5/27                              | 6/26                             |
| Mean age (SD)            | 20.97 (2.47)                      | 20.97 (4.04)                      | 21.47 (2.42)                     |
| Prior knowledge EMDR (%) | 53.13%                            | 53.13%                            | 62.5%                            |

On the pre-screening questionnaire, individuals indicated whether they had prior knowledge about three different therapies: Applied Relaxation, EMDR therapy, and Mindfulness. This was done in order to not draw attention to this focus of this study. There were no differences between the three groups regarding knowledge of EMDR, age, or sex distribution



( $ps > .5$ ; see Table 1). The protocol of this study was approved by the ethics committee of the Faculty of Social and Behavioral Sciences of Utrecht University (code: amendment to 15-080)

### **Materials and Procedure**

The experiment took place in a laboratory at Utrecht University. Upon arriving in the laboratory, participants were again asked about their prior knowledge of three different therapies (Applied Relaxation, EMDR therapy, and Mindfulness) and to rate to what extent they thought that each of these is an effective therapy on a 10-point scale ranging from 1 (*not at all effective*) to 10 (*very effective*). Afterwards, participants were provided with an information letter about the experiment which the experimenter read aloud to them. Depending on the condition, they received information that was positive, negative, or neutral about the effectiveness of EMDR therapy. Participants in the positive information condition were told: “*The experiment you are about to take part in is based on a treatment that focuses on the processing of traumatic or unpleasant memories. This method is called Eye-Movement Desensitization Reprocessing (EMDR) therapy. With this method, it is assumed that carrying out eye-movements while retrieving the memory makes the unpleasant memory less vivid and emotional, so that patients suffer less from this memory. In the past decades, a lot of research has been done into this treatment and the results are very positive. EMDR therapy is a very effective method to reduce the emotionality and vividness of unpleasant memories. This makes it a popular method for the treatment of post-traumatic stress disorder and anxiety symptoms.*” Participants in the negative information condition received the same general introduction about EMDR therapy, but were then told: “*In the past decades, much research has been done into EMDR therapy, from which various results emerge. A small number of studies give moderately positive results, but most researchers conclude that EMDR therapy is not effective in making memories less vivid and*

*emotional. Many researchers are therefore skeptical about the effectiveness of EMDR therapy, despite the popularity of the method.*” Finally, in the neutral information condition also received the same general introduction to EMDR therapy, but were told: “*Little research has been done into the effectiveness of this method.*” The information letter ended with the following instructions: “*In the current experiment we will measure whether making eye-movements affects unpleasant memories that you have about moments from your own life.*”

After the instructions manipulation, participants continued with the eye-movements intervention phase. Following the procedure developed by van den Hout et al (2001), and commonly used in similar studies (Gunter & Bodner, 2008; Littel et al., 2017; Mertens et al., 2018; van Veen et al., 2020), participants were first instructed to recall two unpleasant and vivid autobiographical memories and rate their unpleasantness on a scale from 0 to 100. Memories rated between 60 and 90 were selected and were then ranked in order of unpleasantness. Subsequently, the participant verbally recounted these memories. Randomization was used to decide which memory would be discussed first. In line with the Dutch EMDR therapy protocol (de Jongh & Ten Broeke, 2012), the participants were instructed to ‘play’ these memories in their minds as vividly as possible and tell the experimenter what they remembered. Then, they were asked to take a ‘screenshot’ in their mind of the most emotionally intense and unpleasant moment. Participants labelled each memory image with a keyword, which was used to refer to the image in the remainder of the experiment.

Participants were then subjected to the laboratory analogue of the EMDR therapy protocol. Participants were seated approximately 60 cm in front of a computer screen which was adjusted to their height. They were instructed to recall one of their aversive memories.

Meanwhile they had to make horizontal eye-movements (EM + Recall) induced by tracking a horizontally moving white dot (approximately 1 cm in diameter) on the screen (speed: 1 left-right-left cycle per second; see Littel et al., 2017), or to watch a black screen without a dot (Recall Only). The order in which they performed these tasks was counterbalanced. Each intervention was completed in six blocks of 24 seconds separated by breaks of 10 seconds (van Schie et al., 2016). Before (pre-test) and after (post-test) each intervention participants recalled the aversive memory for 10 seconds and reported its emotionality and vividness ratings using Visual Analog Scales (VASs) ranging from 0 (*not unpleasant/not vivid*) to 100 (*very unpleasant/very vivid*). After finishing both tasks, participants completed a questionnaire, inquiring about the effort the participant invested into the EM + Recall (1-5 scale; 1 = “did not invest effort at all”, 5 = “invested a lot of effort”), whether they believe EMDR therapy is an effective treatment (1-10 scale; 1 = “not at all effective”, 10 = “very effective”), and whether they thought this study attempted to influence this belief (options: “yes”, “no”, “uncertain”). The experimental task was programmed using Inquisit (v4; [www.millisecond.com](http://www.millisecond.com)).

After the computer task, participants were invited to participate in an online follow-up test approximately 24h later presented using Qualtrics (<https://www.qualtrics.com>). They were told that they would receive an additional €2 if they completed this questionnaire. Eighty-four (87.5%) participants did this. In the follow-up test, participants were instructed to recall the aversive memories for 10 seconds and report their emotionality and vividness ratings on VASs similar to the pre and post-test. Afterwards, they were debriefed.

## Results

### *Expectations Regarding EMDR Effectiveness*

Participants' EMDR effectiveness ratings before and after the experiment were analyzed using a repeated measures ANOVA with factors Time (pre- versus post-measurement; within-subjects factor) and Condition (positive, negative, and neutral information; between-subjects factor). This analysis only revealed a significant effect for Time,  $F(1, 91) = 11.59, p = .001, \eta^2_p = .11$ , showing an overall decrease in effectiveness ratings (see Table 2). There were no effects for Condition,  $F(1, 91) = 2.27, p = .109, \eta^2_p = .05$ , or for Time  $\times$  Condition,  $F(1, 91) = 1.50, p = .228, \eta^2_p = .03$ .

Table 2.

*Mean (standard deviation) EMDR therapy effectiveness ratings (1-10) before (pre-rating) and after (post-rating) the study was completed.*

|             | Positive information condition | Negative information condition | Neutral information condition |
|-------------|--------------------------------|--------------------------------|-------------------------------|
| Pre-rating  | 7.16 (1.49)                    | 6.87 (1.23)                    | 7.25 (1.50)                   |
| Post-rating | 6.90 (1.30)                    | 6.03 (1.49)                    | 6.88 (1.66)                   |

### ***Vividness and Emotionality of the Autobiographical Memories***

*Pre-test versus post-test comparison.* Ratings of memory' vividness and emotionality were analyzed with two repeated measures ANOVAs with factor Time (pre-rating, post-ratings; within-subjects factor), Task (EM+Recall, Recall Only; within-subjects factor) and Condition (positive, negative, and neutral information; between-subjects factor).

For vividness, a main effect of Time,  $F(1, 93) = 13.64, p < .001, \eta^2_p = .13$ , and Task,  $F(1, 93) = 7.10, p = .009, \eta^2_p = .07$ , as well as a two-way interaction between these two factors,  $F(1, 93) = 13.30, p < .001, \eta^2_p = .13$ , were found. This interaction reflects a greater decrease from the pre-test to post-test in the EM+Recall compared to Recall Only condition (see Table 3).

However, no main effect or interaction effects were found with Condition,  $F_s < 2.71$ ,  $p_s > .07$ ,  $\eta^2_p < .06$ .

Analysis of the emotionality ratings likewise showed a main effect of Time,  $F(1, 93) = 41.64$ ,  $p < .001$ ,  $\eta^2_p = .31$ , and an interaction between Time and Task,  $F(1, 93) = 13.33$ ,  $p < .001$ ,  $\eta^2_p = .13$ , also reflecting a greater decrease from the pre to post-test in the EM + Recall compared to Recall Only condition (see Table 3). However, also for emotionality ratings, no main effect or interaction effects were found with Condition, all  $F's < 1.31$ , all  $p's > .27$ ,  $\eta^2_p < .03$ .

*Pre-test versus follow-up comparison.* Ratings of memory' vividness and emotionality from the follow-up test were analyzed with two repeated measures ANOVA's with factor Time (pre-rating, follow-up; within-subjects factor), Task (EM+Recall, Recall Only; within-subjects factor) and Condition (positive, negative, and neutral information; between-subjects factor).

For vividness ratings, only a main effect of Time was found,  $F(1, 81) = 80.94$ ,  $p < .001$ ,  $\eta^2_p = .50$ , indicating a large reduction in rated vividness from pre-test to follow-up test (see Table 3). All other main or interaction effects were not statistically significant, all  $F's < 2.53$ , all  $p's > .11$ ,  $\eta^2_p < .04$ .

For emotionality ratings, only a main effect of Time was found,  $F(1, 81) = 63.16$ ,  $p < .001$ ,  $\eta^2_p = .44$ , similarly indicating a large reduction in rated emotionality from pre-test to follow-up test (see Table 3). All other main or interaction effects were not statistically significant, all  $F's < 2.81$ , all  $p's > .06$ ,  $\eta^2_p < .07$ .

Table 3.

*Mean (SD) of vividness and emotionality ratings in the different conditions of Experiment 1.*

|                     | Positive information condition |               | Negative information condition |               | Neutral information condition |               |
|---------------------|--------------------------------|---------------|--------------------------------|---------------|-------------------------------|---------------|
|                     | EM+Recall                      | Recall Only   | EM+Recall                      | Recall Only   | EM+Recall                     | Recall Only   |
| <b>Vividness</b>    |                                |               |                                |               |                               |               |
| Pre-ratings         | 75.90 (9.88)                   | 77.21 (11.15) | 74.73 (16.50)                  | 72.62 (16.98) | 80.52 (12.40)                 | 81.69 (10.43) |
| Post-ratings        | 66.17 (20.87)                  | 74.69 (15.87) | 66.35 (21.56)                  | 74.12 (15.74) | 73.76 (19.75)                 | 79.07 (14.16) |
| Follow-up           | 64.03 (13.80)                  | 62.90 (15.06) | 62.08 (16.98)                  | 58.15 (19.01) | 65.28 (19.23)                 | 62.83 (20.38) |
| <b>Emotionality</b> |                                |               |                                |               |                               |               |
| Pre-ratings         | 70.24 (12.49)                  | 70.00 (11.03) | 73.04 (16.98)                  | 71.81 (13.25) | 75.10 (14.10)                 | 72.38 (10.38) |
| Post-ratings        | 61.10 (17.91)                  | 63.79 (17.81) | 66.19 (19.23)                  | 70.42 (14.57) | 63.69 (15.99)                 | 69.34 (13.08) |
| Follow-up           | 57.45 (20.37)                  | 60.62 (13.33) | 62.81 (15.78)                  | 58.31 (14.44) | 58.48 (18.10)                 | 60.97 (16.60) |

### ***Correlation Analyses***

Because our manipulation of EMDR effectiveness ratings did not differ per condition, we additionally used a different approach to investigate whether EMDR's effectiveness expectations were related to the effectiveness of the intervention. Therefore, we correlated effectiveness ratings before the experiment with the pre-to-posttest and pre-to-follow-up test differences between the EM+Recall condition and the Recall Only condition. These correlations were not significant for both the pre-to-posttest difference (vividness:  $r = .059$ ,  $p = .570$ ; emotionality:  $r = -.048$ ,  $p = .644$ ) and the pre-to-follow-up test difference (vividness:  $r = .031$ ,  $p = .780$ ; emotionality:  $r = -.101$ ,  $p = .635$ ).

### **Discussion**

In this first study, we found no evidence that a laboratory analog of EMDR therapy is influenced by participants' treatment effectiveness expectations. However, one important

limitation is that we were unable to successfully manipulate participants' expectations. This may be because many participants were familiar with EMDR therapy and it may be difficult to change expectations for an existing and (in the Netherlands) well-known therapy. Therefore, we decided to conduct a second study using a novel task that capitalizes on the same working mechanisms as making eye-movements in EMDR, but did not overtly resemble EMDR therapy. This way we reduced the possibility that our results are affected by participants' prior knowledge about and familiarity with EMDR therapy.

## **Experiment 2**

Experiment 2 was an adapted and extended version of Experiment 1. We made several changes compared to Experiment 1. First, we used a different intervention to induce eye-movements dubbed the "Working Memory-Symbol Recognition Task (WM-SRT)" to mask any connection with EMDR therapy. In this task participants see distractor letters (e.g., "m") appear alternatingly on the left and right side of the computer screen. Participants are instructed to press the spacebar whenever they see a target letter (e.g., "n"). Participants are also instructed to keep their head still and only move their eyes. Hence, this task induces left-right eye-movements (Homer, Deeprose, and Andrade, 2016). A second change is that the study was conducted online to facilitate testing a larger sample (i.e., more statistical power). The third change was a more extensive assessment of participants' expectations using the Expectations/Credibility Questionnaire (CEQ; Devilly and Borkovec, 2000) at the end of the study. Finally, we removed the neutral information condition to focus on the two most extreme conditions and to further increase the statistical power.

## **Pre-registration**

The sample size determination, design, procedure, and data analyses steps were pre-registered on the Open Science Framework prior to the data collection (<https://osf.io/mu3ca/>).

## **Participants**

Based on an a priori power analysis, we decided to test 80 participants in each condition (160 in total). This sample size provides adequate statistical power ( $> 0.80$ ) to detect small-to-medium sized differences (Cohen's  $d = .4$ ) between the two groups with an alpha cut-off of .05. Participants were recruited through posters at the Utrecht University and Erasmus University Rotterdam campus and online advertisement. Two-hundred and thirty-eight participants clicked on the participation link. Based on our exclusion criteria (currently being under treatment by a psychologist/psychiatrist or reporting complaints regarding unpleasant and/or intrusive memories), 49 participants were automatically excluded based on an initial screening. The remaining 189 participants provided informed consent to participate in the study. Of these participants, 178 completed the full experiment and answered the post-rating assessment. Finally, five participants were removed because they failed to select a sufficiently unpleasant memory (rated less than 5 on a 1-10 scale;  $n = 3$ ), indicated that they did not pay sufficient attention during the experiment (rated less than 5 on a 1-10 scale;  $n = 1$ ), or completed the experiment twice ( $n = 1$ ). A sample of 173 participants was included for the final analyses. These participants were randomly assigned to the two conditions: positive information ( $n = 83$ ; mean age = 22.76, SD = 5.79; 67 women, 16 men) and negative information ( $n = 90$ ; mean age = 22.26, SD = 6.84; 72 women, 16 men, 2 preferred not to say). The groups did not differ in age or sex ( $ps > 0.6$ ). Participants were rewarded with study credit or a chance (1/10) to win a coupon (€10). The



protocol of this study was approved by the ethics committee of the Faculty of Social and Behavioral Sciences of Utrecht University (code: 19-122).

## Materials and Procedure

The task was programmed in Inquisit v4 (<https://www.millisecond.com/>) and administered online through the Inquisit Web servers. Participants were asked to execute the task from home in a quiet surrounding using their laptop or personal computer. The task started with a pre-screening assessment, an information letter, and an informed consent form. The information letter was manipulated to be either very positive or very negative about a novel treatment called “Working Memory-Symbol Recognition Task”. Particularly, participants in the positive information condition were told that: *“The experiment in which you are about to participate investigates a treatment that focuses on the processing of unpleasant or traumatic memories. This relatively new treatment is based on focusing the attention to symbols while thinking back of the memory. This treatment is called the Working Memory Symbol recognition Task (WM-SRT). In this treatment, it is presupposed that loading the working memory while recollection a memory will make this memory less vivid and emotional, thereby reducing the distress of the memory for patients. In the past decennia, much research has been conducted on WM-SRT, and most studies have concluded that WM-SRT is very effective to make memories less vivid and emotional. Many researchers and clinicians are therefore extremely enthusiastic about the effectiveness of WM-SRT.”* Participants in the negative information condition were given the same general introduction about WM-SRT, but were additionally informed that: *“In the past decennia, much research has been conducted on WM-SRT, in which different results have been found. A small number of studies indicated moderate positive results, but the majority of studies*

*have concluded that WM-SRT is not an effective treatment to make memories less vivid and emotional. Many researchers and clinicians are therefore skeptical about the effectiveness of WM-SRT.”* The information letter ended with the following information: “*In this study we will measure whether loading working memory has an effect on unpleasant memories you have about certain moments in your life.*” Following this manipulation, participants were asked to select one unpleasant memory about a situation in the past (> 1 week ago) that, when recalled, would still made them feel emotional. They were asked to provide one keyword that would trigger them to think about the memory.

After the expectancy manipulation and memory selection, participants rated memory vividness and emotionality following the same procedure as Experiment 1, with the addition that they also scored their memory accessibility on a 0-100 VAS (“How easily could you recollect the memory?”; 0 = not at all easily, 100 = very easily). Next, they completed four blocks of 24s of the task developed by Homer et al. (2016). Before each block, participants were asked to think back to the selected memory (using the keyword they selected) and to keep it in mind while executing the task. Furthermore, they were told what the target letter was (i.e., ‘n’, ‘p’, ‘v’, or ‘e’) and they were instructed to press the space bar whenever they saw this target letter appear. The task consisted of 300 ms presentations of letters on the left and right sides of the computer screen (with a 450 ms inter-trial interval). Each block consisted of 30 distractor letter presentations (i.e., ‘m’, ‘d’, ‘w’, or ‘c’) and two target letter presentations. A different target and distractor letter was used in each block. In the background, alternating black and white stripes were presented to increase the visual load of the task (see Homer et al., 2016). A prior study from our lab showed that this task was comparably effective to a dot-tracking task to reduce the

reported emotionality and vividness of negative autobiographical memories (Mertens et al., 2018). After completing the task, participants provided post-ratings of their selected memory in the same way as in the pre-test.

At the end of the experiment, participants completed an adjusted version of the Credibility/Expectancy Questionnaire (CEQ; Devilly & Borkovec, 2000). This adjusted CEQ included only the first four questions of this questionnaire and questions were rephrased to refer specifically to the WM-SRT therapy.<sup>1</sup> Finally, participants were asked to indicate to what extent they paid attention to the task (0-10 scale; 0 = “not at all attentive”, 5 = “somewhat attentive”, 10 = “very attentive”) with the explicit instruction that their answer to this question would not impact their compensation for their participation. The experiment took approximately 20 minutes.

## **Data reduction and analysis**

### ***Expectations Regarding EMDR effectiveness (Credibility/Expectations Questionnaire)***

First, the psychometric properties of the CEQ were evaluated. A principle components analysis confirmed that all questions loaded acceptably onto one common latent factor (63.45% explained variance) and that internal consistency was good (Cronbach's alpha = 0.80). Therefore, a weighted sum score of the scale was calculated and analyzed using a one-way ANOVA with Condition (positive information, negative information) as a between-subjects

---

<sup>1</sup> This was decided because the CEQ was developed to be administered before (four questions) and after (two questions) psychotherapy. Therefore, the original questions in the CEQ do not refer to a specific therapy, but only to ‘the therapy’. For use in our study, we decided to reword the questions to more clearly refer to the ‘WM-SRT’ intervention and only use the first four question since we only had one measurement time point.

factor. Although descriptively the results were in the expected direction, no statistically significant difference was found between the two conditions for the CEQ ( $p = .114$ , Cohen's  $d = 0.24$ ; see Table 4).

Table 4.

*Credibility/expectancy questionnaire results for the two conditions in Experiment 2*

| CEQ items   | Mean ( <i>SD</i> )<br>positive<br>condition | Mean ( <i>SD</i> )<br>negative<br>condition | Comparison           |
|---|---|---|----------------------|
| 1. How logical did you find the WM-SRT therapy? (1-9)   | 5.72 (2.08)                                 | 5.22 (1.87)                                 | $F(1, 168) = 2.76^a$ |
| 2. How successful do you think the WM-SRT was to reduce the intensity of your memories? (1-9)                   | 5.13 (1.78)                                 | 4.82 (1.87)                                 | $F(1, 168) = 1.27$   |
| 3. How much confidence would you have to recommend this therapy to a friend? (1-9)                              | 5.16 (2.02)                                 | 4.61 (1.78)                                 | $F(1, 168) = 3.50^b$ |
| 4. At the end of the task, how strongly did you think the WM-SRT reduced the intensity of your memory? (0-100%) | 28.90%<br>(17.99%)                          | 28.07%<br>(21.38%)                          | $F(1, 168) = 0.08$   |
| Weighted mean   | 18.61 (5.79)                                | 17.17 (6.01)                                | $F(1, 168) = 2.52$   |

*Note:* <sup>a</sup> $p = .099$ ; <sup>b</sup> $p = .075$

### ***Vividness, Emotionality, and Accessibility of the Autobiographical Memories***

All three ratings were analyzed using a repeated measure ANOVA with factor Time (pre-ratings, post-ratings; within-subjects) and Condition (positive information, negative information; between-subjects). For all outcome measures (vividness, emotionality, and accessibility) there was a clear effect of Time: vividness,  $F(1, 171) = 42.23$ ,  $p < .001$ ,  $\eta^2_p = .20$ , emotionality,  $F(1, 171) = 49.98$ ,  $p < .001$ ,  $\eta^2_p = .23$ ; and accessibility,  $F(1, 171) = 64.00$ ,  $p < .001$ ,  $\eta^2_p = .27$ . This indicates that the intervention reduced the memory across these three dimensions (see Table 5).

However, for neither outcome measure, an interaction with factor Condition was found:

vividness:  $F(1, 171) = 0.60, p = .439, \eta^2_p < .01$ ; emotionality:  $F(1, 171) = 0.62, p = .431, \eta^2_p = .00$ ; and accessibility:  $F(1, 171) = 0.14, p = .711, \eta^2_p = .00$ ). This result indicates that that our manipulation did not modulate the effect of the intervention.

Table 5.

*Mean (SD) memory vividness, emotionality and accessibility rating across the different conditions of Experiment 2.*

|                      | Positive information condition<br>( $n = 83$ ) | Negative information condition<br>( $n = 90$ ) |
|----------------------|--|--|
| <b>Vividness</b>     |  |  |
| Pre-test             | 72.17 (21.79)                                  | 70.78 (22.29)                                  |
| Post-test            | 60.46 (21.94)                                  | 61.57 (21.90)                                  |
| <b>Emotionality</b>  |  |  |
| Pre-test             | 69.52 (15.71)                                  | 68.48 (16.59)                                  |
| Post-test            | 62.60 (18.34)                                  | 59.82 (17.44)                                  |
| <b>Accessibility</b> |  |  |
| Pre-test             | 76.49 (20.78)                                  | 75.31 (19.99)                                  |
| Post-test            | 63.25 (22.00)                                  | 63.24 (23.31)                                  |

## Discussion

Experiment 2 was an online study, in which we attempted to manipulate participants' expectations regarding a non-existing treatment capitalizing on the same mechanisms as EMDR (i.e., WM-SRT therapy). Despite a well-powered study ( $N = 173$ ), we were unable to find a significant impact of our treatment effectiveness instructions on participants' expectations of the treatment as measured with the CEQ. Furthermore, we found a clear reduction in memory vividness, emotionality, and accessibility due to the intervention, but no interaction with our manipulation. Taken together, this second experiment confirmed the results of the first study: participants' treatment expectations did not modulate the memory degrading effects of the eye-movements intervention. However, as in Experiment 1, expectations of the effectiveness of the

intervention did not significantly differ between groups receiving positive and negative information. Hence the absence of effects of the effectiveness expectations manipulation on the outcome measures (i.e., memory vividness, emotionality and accessibility) is difficult to interpret due to the absence of a successful manipulation.

### **General Discussion**

Despite ongoing debate about the role of non-specific factors in psychotherapy in general (Cuijpers et al., 2019) and EMDR therapy in particular (Deville, 2005; Lilienfeld, 1996), there is a lack of research directly addressing the role of such factors. In two pre-registered studies, we investigated the role of treatment effectiveness expectations in a laboratory model of EMDR therapy. We found no evidence that verbal suggestions about treatment expectations impacted the effectiveness of our laboratory models of EMDR therapy. These results corroborate the only two other available reports on this topic (Gosselin & Matthews, 1995; Littel et al., 2017) and suggest that that treatment expectations do not modulate the effectiveness of a laboratory analogue of EMDR therapy.

For clinical purposes, our results are encouraging, because they indicate that a laboratory analogue of EMDR therapy seems to work, regardless of (negative) verbal suggestions about its effectiveness. Hence, EMDR therapy, and potentially other psychotherapies, may work because of specific factors present in the protocols, rather than non-specific factors present in most psychotherapies such as verbal suggestions about the effectiveness of the therapy. This provides support for the continued development of psychotherapy protocols according to mechanistic theories about human psychopathology (e.g., Craske et al., 2014) and training aspiring psychotherapists in these protocols.

One important observation in our studies is that it was difficult to significantly change participants' treatment effectiveness expectations using verbal suggestions. This was both the case for EMDR therapy (for which many of our participants had pre-existing knowledge), but also for a non-existing therapy for which participants could not have had any prior knowledge. This indicates that verbal suggestions appear to be a relatively weak route to change therapy effectiveness expectations. However, some nuance is in place here. First, while in the first experiment the information letter and thus also the manipulation were read aloud to the participants, in the second experiment the manipulation was only delivered via the informed consent letter. Though participants were asked to carefully read this letter, there is evidence that participants in psychology studies often do not carefully attend to the information presented to them (e.g., Douglas et al., 2020). This may have limited the effectiveness of our manipulation, particularly in the second experiment. Furthermore, the letter was delivered within the setting of experimental research in university students. It is likely that students are quite skeptical and critical concerning information provided to them within this setting (Fransen et al., 2015). Hence, it may be that through other mediums (e.g., direct instructions rather than in an information letter), in different settings (e.g., within a clinic) and in other populations (e.g., patients; general population), it may be the case the verbal suggestions are more effective to change expectations regarding therapy effectiveness (see Testa & Rossetini, 2016).

Another aspect of our results that merits some discussion is the fact that our interventions were quite effective in changing the reported emotionality and vividness of autobiographical memories, even in an online-based implementation of our task. The observed reductions observed in both studies (approximately 10 points on a 100 scale for all outcome measures) is

comparable to the average reduction found in the eye-movements condition in previous studies (Homer et al., 2016; van Schie et al., 2016; van Veen et al., 2020). Hence, an online implementation of the EMDR-analog paradigm appears feasible, which opens up opportunities for future research.

A number of limitations of these studies can be noted. First, our research was conducted with a student sample in a non-clinical setting. This limits the extent to which our results can be generalized to other populations and different settings. Nonetheless, this type of research can be challenging, because inducing negative expectations about treatment in patients may adversely impact their treatment willingness and treatment success. Hence, though research within clinical contexts is optimal, fundamental research is needed investigate the impact of treatment effectiveness expectations in a safe and well-controlled environment. Another limitation is that we were unable to reliably change treatment effectiveness expectations. This lack of a successful manipulation complicates the interpretation of our results. However, we believe that these two relative large experiments provide an important addition to the literature regarding the role of treatment expectation in EMDR therapy.

Taken together, in two laboratory analogues of EMDR therapy, clear reductions in memory vividness and emotionality were found, despite different suggestions regarding the effectivity of the intervention. This suggests that treatment effectiveness expectancy is perhaps not part of the ‘active ingredients’ of EMDR therapy. Nonetheless, more research is needed, particularly focusing on stronger treatment expectation manipulations and including patients in clinical settings.



### References

- Andrade, J., Kavanagh, D., & Baddeley, A. (1997). Eye-movements and visual imagery: A working memory approach to the treatment of post-traumatic stress disorder. *British Journal of Clinical Psychology*, 36(2), 209–223. <https://doi.org/10.1111/j.2044-8260.1997.tb01408.x>
- Arnkoff, D. B., Glass, C. R., & Shapiro, J. S. (2002). Expectations and preferences. In J. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 335–356). Oxford University Press.
- Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, 63(1), 1–29. <https://doi.org/10.1146/annurev-psych-120710-100422>
- Constantino, M. J., Arnkoff, D. B., Glass, C. R., Ametrano, R. M., & Smith, J. Z. (2011). Expectations. *Journal of Clinical Psychology*, 67(2), 184–192. <https://doi.org/10.1002/jclp.20754>
- Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, 58, 10–23. <https://doi.org/10.1016/j.brat.2014.04.006>
- Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). The Role of Common Factors in Psychotherapy Outcomes. *Annual Review of Clinical Psychology*, 15(1), 207–231. <https://doi.org/10.1146/annurev-clinpsy-050718-095424>
- de Jongh, A., & Ten Broeke, E. (2012). *Handboek EMDR: Een geprotocolleerde behandelmethode voor de gevolgen van psychotrauma*. Pearson.
- Devilly, G. J. (2005). Power Therapies and possible threats to the science of psychology and psychiatry. *Australian and New Zealand Journal of Psychiatry*, 39(6), 437–445.

<https://doi.org/10.1111/j.1440-1614.2005.01601.x>

Deville, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 31(2), 73–86.

[https://doi.org/10.1016/S0005-7916\(00\)00012-4](https://doi.org/10.1016/S0005-7916(00)00012-4)

Douglas, B., McGorray, E., & Ewell, P. (2020). Some researchers wear yellow pants, but even fewer participants read consent forms: Exploring and improving consent form reading in human subjects research. *Preprint*. <https://doi.org/10.31234/osf.io/hcvk6>

Engelhard, I. M., McNally, R. J., & van Schie, K. (2019). Retrieving and Modifying Traumatic Memories: Recent Research Relevant to Three Controversies. *Current Directions in Psychological Science*, 12(1), 096372141880772.

<https://doi.org/10.1177/0963721418807728>

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>

Fransen, M. L., Smit, E. G., & Verlegh, P. W. J. (2015). Strategies and motives for resistance to persuasion: an integrative framework. *Frontiers in Psychology*, 6.

<https://doi.org/10.3389/fpsyg.2015.01201>

Gosselin, P., & Matthews, W. J. (1995). Eye movement desensitization and reprocessing in the treatment of test anxiety: A study of the effects of expectancy and eye movement. *Journal of Behavior Therapy and Experimental Psychiatry*, 26(4), 331–337.

[https://doi.org/10.1016/0005-7916\(95\)00038-0](https://doi.org/10.1016/0005-7916(95)00038-0)

Greenberg, R. P., Constantino, M. J., & Bruce, N. (2006). Are patient expectations still relevant for psychotherapy process and outcome? *Clinical Psychology Review*, 26(6), 657–678.

<https://doi.org/10.1016/j.cpr.2005.03.002>

Gunter, R. W., & Bodner, G. E. (2008). How eye movements affect unpleasant memories:

Support for a working-memory account. *Behaviour Research and Therapy*, 46(8), 913–931.

<https://doi.org/10.1016/j.brat.2008.04.006>

Homer, S. R., Deeprose, C., & Andrade, J. (2016). Negative mental imagery in public speaking anxiety: Forming cognitive resistance by taxing visuospatial working memory. *Journal of Behavior Therapy and Experimental Psychiatry*, 50, 77–82.

<https://doi.org/10.1016/j.jbtep.2015.05.004>

Lewis, C., Roberts, N. P., Andrew, M., Starling, E., & Bisson, J. I. (2020). Psychological therapies for post-traumatic stress disorder in adults: systematic review and meta-analysis.

*European Journal of Psychotraumatology*, 11(1), 1729633.

<https://doi.org/10.1080/20008198.2020.1729633>

Lilienfeld, S. O. (1996). EMDR treatment: Less than what meets the eye? *Skeptical Inquirer*, 25–31. <https://www.quackwatch.org/01QuackeryRelatedTopics/emdr.html>

Littel, M., van Schie, K., & van den Hout, M. A. (2017). Exploring expectation effects in

EMDR: does prior treatment knowledge affect the degrading effects of eye movements on memories? *European Journal of Psychotraumatology*, 8(sup1), 1328954.

<https://doi.org/10.1080/20008198.2017.1328954>

Mertens, G., Krypotos, A.-M., van Logtestijn, A., Landkroon, E., Veen, S. C. van, & Engelhard,

I. M. (2018). Changing negative autobiographical memories in the lab: a comparison of three eye-movement tasks. *Memory*, 1–11. <https://doi.org/10.1080/09658211.2018.1507041>

Shapiro, F. (2017). *Eye movement desensitization and reprocessing (EMDR) therapy: Basic principles, protocols, and procedures* (3rd ed.). Guilford Press.

Testa, M., & Rossetini, G. (2016). Enhance placebo, avoid nocebo: How contextual factors affect physiotherapy outcomes. *Manual Therapy*, 24, 65–74.

<https://doi.org/10.1016/j.math.2016.04.006>

van den Hout, M. A., & Engelhard, I. M. (2012). How does EMDR work? *Journal of Experimental Psychopathology*, 3(5), jep.028212. <https://doi.org/10.5127/jep.028212>

van Schie, K., van Veen, S. C., Engelhard, I. M., Klugkist, I., & van den Hout, M. A. (2016). Blurring emotional memories using eye movements: individual differences and speed of eye movements. *European Journal of Psychotraumatology*, 7(1), 29476.

<https://doi.org/10.3402/ejpt.v7.29476>

van Veen, S. C., van Schie, K., van de Schoot, R., van den Hout, M. A., & Engelhard, I. M. (2020). Making eye movements during imaginal exposure leads to short-lived memory effects compared to imaginal exposure alone. *Journal of Behavior Therapy and Experimental Psychiatry*, 67, 101466. <https://doi.org/10.1016/j.jbtep.2019.03.001>